JOURNAL OF MULTIVARIATE ANALYSIS 43, 1–15 (1992)

Estimation of Mean Square Error of Empirical Best Linear Unbiased Predictors under a Random Error Variance Linear Model

J. KLEFFE

University of Copenhagen, Copenhagen, Denmark

AND

J. N. K. RAO

Carleton University, Ottawa, Canada

Communicated by the Editors

A linear model with random effects, μ_i, and random error variances, σ_i, is considered. The linear Bayes estimator or the best linear unbiased predictor (BLUP) of μ_i is first obtained, and then the unknown parameters in the model are estimated to arrive at the empirical linear Bayes estimator or the empirical BLUP (EBLUP) of μ_i. A second-order approximation to mean square error (MSE) of the EBLUP and an approximately unbiased estimator of MSE are derived. Results of a simulation study confirm the accuracy of these approximations.

1. INTRODUCTION

Consider the following linear model with random error variances:

$$y_{ij} = \mu_i + e_{ij}, \qquad i = 1, ..., m \qquad j = 1, ..., n, \tag{1.1}$$

with

$$\mu_i \sim N(\mu, \tau), \qquad (e_{ij}|\sigma_i) \sim N(0, \sigma_i), \qquad \sigma_i \sim (\beta, \alpha), \tag{1.2}$$

where μ_i and e_{ij} are all independently distributed given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$. Further, the error variances σ_i are assumed to be nonnegative i.i.d. random variables with mean β and variance α and independent of μ_i and e_{ij}. The special case of equal error variances, $\sigma_i = \beta$, has received considerable

Received March 26, 1991; revised January 13, 1992.

AMS 1980 subject classifications: primary 62J99, secondary 62D05.

Key words and phrases: estimation of random effects, second-order approximation to MSE, small area estimation.

0047-259X/92 $5.00

attention in the literature. In particular, the linear Bayes estimator or the best linear unbiased predictor (BLUP) of μ_i is first obtained, and then the unknown parameters in the model are estimated to arrive at the empirical Bayes estimator or the empirical BLUP (EBLUP) of μ_i (see, e.g., [2]). Prasad and Rao [3] derived second-order approximations to the mean square error (MSE) of the EBLUP and the estimator of MSE, assuming equal error variances and large m.

The main purpose of this article is to obtain a second order-approximation to MSE of the EBLUP of μ_i and an approximately unbiased estimator of the MSE under the general model (1.1) with random error variances. These approximations are correct up to terms of order $1/m$. Finite sample properties are also studied through a simulation study. Aragon [1] studied more general models with random error variances, but he did not consider the prediction of random effects like μ_i. Instead, he considered the estimation of parameters like μ, τ, β, and α, assuming inverse Gaussian errors with variances σ_i.

The results derived in Sections 2, 3, and 4 should be useful in small area estimation, where μ_i and σ_i correspond to ith small area mean and variance. It is more realistic to assume random small area variances than a constant variance across small areas.

2. Derivation of EBLUP

Let $y_{i\cdot} = \sum_j y_{ij}/n$, $y_{\cdot\cdot} = \sum_i y_{i\cdot}/m$, and $\delta = \tau + \beta/n$. Then it is easy to see that the BLUP (or the linear Bayes estimator) of μ_i under the general model (1.1) is given by

$$\tilde{\mu}_i = y_{i\cdot} - c(y_{i\cdot} - y_{\cdot\cdot}), \qquad c = \beta/n\delta. \tag{2.1}$$

It is interesting to note that the same predictor is obtained under the assumption of equal error variances, $\sigma_i = \beta$.

The BLUP $\tilde{\mu}_i$ depends on unknown parameters β and δ. We use simple moments estimates of β and δ to obtain an EBLUP of μ_i. An obvious unbiased estimator of δ is

$$\hat{\delta} = (m-1)^{-1} \sum_{i=1}^{m} (y_{i\cdot} - y_{\cdot\cdot})^2 \tag{2.2}$$

with conditional expectation $E_2\hat{\delta} = \tau + \sigma_{\cdot}/n$ given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$, where $\sigma_{\cdot} = \sum_i \sigma_i/m$. For the parameter β, we use the unbiased estimator

$$\hat{\beta} = \mathrm{m}^{-1}(n-1)^{-1} \sum_{i=1}^{m} \sum_{j=1}^{n} (y_{ij} - y_{i\cdot})^2 = \sum_{i=1}^{m} \hat{\sigma}_i/m \tag{2.3}$$

which satisfies $E_2\hat{\beta}=\sigma..$ The estimators $\hat{\beta}$ and $\hat{\delta}$ are conditionally independent, since $\hat{\beta}$ depends only on the within groups sum of squares while $\hat{\delta}$ is a function of the within group means.

Substituting $\hat{\beta}$ and $\hat{\delta}$ for β and δ in (2.1) we obtain the EBLUP as

$$\hat{\mu}_i = y_{i\cdot} - \hat{c}(y_{i\cdot} - y_{\cdot\cdot}), \qquad \hat{c} = \hat{\beta}/n\hat{\delta}. \tag{2.4}$$

Again, the same EBLUP is obtained under the assumption of equal error variances, but the MSE of $\hat{\mu}_i$ under the general model (1.1) will be different.

3. MSE Approximation

3.1. *MSE of the BLUP*

Before deriving a second-order approximation to MSE of the EBLUP, we derive the MSE of the BLUP, $\tilde{\mu}_i$. We write $\tilde{\mu}_i - \mu_i = e_{i\cdot} - c(y_{i\cdot} - y_{\cdot\cdot})$ so that

$$(\tilde{\mu}_i - \mu_i)^2 = e_{i\cdot}^2 - 2c(y_{i\cdot} - y_{\cdot\cdot})e_{i\cdot} + c^2(y_{i\cdot} - y_{\cdot\cdot})^2, \tag{3.1}$$

where $e_{i\cdot} = \sum_j e_{ij}/n$. We next evaluate the conditional expectation, E_2, of the terms in (3.1) given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$.

Using the decomposition $y_{i\cdot} - y_{\cdot\cdot} = (\mu_i - \mu_\cdot) + (e_{i\cdot} - e_{\cdot\cdot})$, where $\mu_\cdot = \sum_i \mu_i/m$ we obtain

$$\begin{aligned} E_2 e_{i\cdot}^2 &= \sigma_i/n \\ E_2(y_{i\cdot} - y_{\cdot\cdot})e_{i\cdot} &= \sigma_i(m-1)/nm \\ E_2(y_{i\cdot} - y_{\cdot\cdot})^2 &= (\tau + \sigma_i/n)(m-1)/m + (\sigma_\cdot - \sigma_i)/nm. \end{aligned} \tag{3.2}$$

Using (3.2) in (3.1) yields the MSE of the BLUP as

$$E(\tilde{\mu}_i - \mu_i)^2 = \tau\beta/n\delta + \beta^2/mn^2\delta. \tag{3.3}$$

It is again interesting to note that the same MSE is obtained under the assumption of equal error variances $\sigma_i = \beta$.

3.2. *Approximation to MSE of the EBLUP*

3.2.1. *A starting Expression for MSE*

Replacing in (3.1) c by its estimator $\hat{c}$, we obtain the squared deviation of $\hat{\mu}_i$ from μ_i, namely $(\hat{\mu}_i - \mu_i)^2$. By symmetry, $E(\hat{\mu}_i - \mu_i)^2$ cannot depend on i, so that by averaging over i and using (2.2), we obtain the MSE of the EBLUP as

$$E(\hat{\mu}_i - \mu_i)^2 = E\left[\sum_i e_{i\cdot}^2 - 2\hat{c}\sum_i (y_{i\cdot} - y_{\cdot\cdot})e_{i\cdot} + (m-1)\,\hat{\beta}\hat{c}/n\right]\Big/ m. \quad (3.4)$$

Since $E_2 \sum_i e_{i\cdot}^2/m = \sigma_{\cdot}/n = E_2\hat{\beta}/n$, an equivalent expression to (3.4) is

$$E(\hat{\mu}_i - \mu_i)^2 = E(\hat{\beta} - \phi_1 + \phi_2)/n,$$

where (3.5)

$$\phi_1 = 2\sum_i (y_{i\cdot} - y_{\cdot\cdot})\, e_{i\cdot}\hat{\beta}/m\hat{\delta}, \qquad \phi_2 = (m-1)\,\hat{\beta}^2/mn\hat{\delta}.$$

Further, we can replace $\hat{\beta}$ and $\hat{\beta}^2$ in (3.5) by $E_2\hat{\beta}$ and $E_2\hat{\beta}^2$, since $\hat{\beta}$ is conditionally independent of $y_{i\cdot}$, $y_{\cdot\cdot}$, $e_{i\cdot}$, and $\hat{\delta}$. Hence, noting that

$$E_2\hat{\beta} = \sigma_{\cdot}, \qquad E_2\hat{\beta}^2 = \sigma_{\cdot}^2 + 2\sum_i \sigma_i^2/(n-1)m^2, \quad (3.6)$$

we obtain

$$E(\hat{\mu}_i - \mu_i)^2 = E(\sigma_{\cdot} - \tilde{\phi}_1 + \tilde{\phi}_2)/n, \quad (3.7)$$

where

$$\tilde{\phi}_1 = 2\sigma_{\cdot} \sum_i (y_{i\cdot} - y_{\cdot\cdot})\, e_{i\cdot}/m\hat{\delta}$$

and

$$\tilde{\phi}_2 = (m-1)\left[\sigma_{\cdot}^2 + 2\sum_i \sigma_i^2/(n-1)m^2\right]/mn\hat{\delta}. \quad (3.8)$$

3.2.2. *Approximations to $E\tilde{\phi}_1$ and $E\tilde{\phi}_2$*

To evaluate (3.7) we need $E\tilde{\phi}_1$ and $E\tilde{\phi}_2$, but no closed-form expressions for these expectations could be obtained. We have therefore derived approximations to $E\tilde{\phi}_1$ and $E\tilde{\phi}_2$ such that the neglected terms are of lower order than $1/m$, for large m. We use the expansion

$$\hat{\delta}^{-1} = [1 - (\hat{\delta} - \delta_1)/\delta_1 + \{(\hat{\delta} - \delta_1)/\delta_1\}^2]/\delta_1 - \{(\hat{\delta} - \delta_1)/\delta_1\}^3/\hat{\delta}, \quad (3.9)$$

where $\delta_1 = E_2\hat{\delta} = \tau + \sigma_{\cdot}/n$. We also need the following two lemmas proved in Section 6.1.

LEMMA 1. *If $m > 4k + 1$ and $\tau > 0$, then $E\hat{\delta}^{-k} \leqslant (2/\tau)^k$, where k is a positive integer.*

LEMMA 2. *Let f be a function with finite second moments, $\tau > 0$ and let σ_i have 4sth moment. Then for $m > 17$,*

$$E[(\hat{\delta} - \delta_1)/\delta_1]^s f/\hat{\delta} = O(m^{-s/2}), \qquad E[(\hat{\delta} - \delta_1)/\delta_1]^s f = O(m^{-s/2}). \tag{3.10}$$

Routine computation based on well-known formulae for second moments of quadratic forms in normal variables shows that

$$E_2(y_{i\cdot} - y_{\cdot\cdot})^2 e_{i\cdot}^2 = \tau\sigma_i(m-1)/mn + 2\sigma_i^2[(m-1)/mn]^2 + \sigma_i[(m-2)\sigma_i + \sigma.]/mn^2. \tag{3.11}$$

This expectation is bounded as a function m. Therefore, assuming finite fourth-order moments for the vector $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$, the functions $f = (y_{i\cdot} - y_{\cdot\cdot})\, e_{i\cdot}\sigma.$ and $g = \sigma_{\cdot}^2 + 2\sum_i \sigma_i^2/(n-1)m^2$ satisfy the assumptions made in Lemma 2. Using (3.9) in (3.8) it follows from (3.10) that

$$E\tilde{\phi}_1 = 2E\left[\sigma.\sum_i (y_{i\cdot} - y_{\cdot\cdot})e_{i\cdot}\right][1 + (\hat{\delta} - \delta_1)^2/\delta_1^2]/m\delta_1 + O(m^{-3/2}) \tag{3.12}$$

and

$$E\tilde{\phi}_2 = (m-1)\, E\left[\sigma_{\cdot}^2 + 2\sum_i \sigma_i^2/(n-1)m^2\right] \times [1 + (\hat{\delta} - \delta_1)^2/\delta_1^2]/mn\delta_1 + O(m^{-3/2}). \tag{3.13}$$

Denote the expectations on the right-hand side of (3.12) and (3.13) by $E\phi_{1*}$ and $E\phi_{2*}$, respectively.

3.2.3. *Conditional Expectation Given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$*

It remains to evaluate $E\phi_{1*}$ and $E\phi_{2*}$. We evaluate these expectations by first evaluating the conditional expectation, E_2, given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$. To evaluate $E_2\phi_{1*}$ we need (see Section 6.2 for details):

$$E_2\left[\sum_i (y_{i\cdot} - y_{\cdot\cdot})e_{i\cdot}\right] = n^{-1}(m-1)\sigma. \tag{3.14}$$

$$E_2\left[\sum_i (y_{i\cdot} - y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}\right] = n^{-1}\left[(m-1)\,\sigma.\delta_1 + 2\sum_i \delta_{1i}\sigma_i/m\right] + O(m^{-1}) \tag{3.15}$$

$$E_2\left[\sum_i (y_{i\cdot} - y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}^2\right] = n^{-1}\left[(m+1)\,\sigma.\delta_1^2 + 2s^2(\tilde{\delta})\sigma. + 4\delta_1\sum_i \delta_{1i}\sigma_i/m\right] + O(m^{-1}). \tag{3.16}$$

Here $\tilde{\delta} = (\delta_{11}, ..., \delta_{1m})'$, $\delta_{1i} = \tau + \sigma_i/n$, $\delta_1 = m^{-1} \sum_i \delta_{1i}$, and $s^2(\tilde{\delta}) = m^{-1} \sum_i (\delta_{1i} - \delta_1)^2$. Using the above expressions in (3.12), we obtain

$$E_2\phi_{1*} = -2\sigma.(mn\delta_1)^{-1}\left[(m+1)\sigma. - 2(m\delta_1)^{-1}\sum_i \delta_{1i}\sigma_i + 2s^2(\tilde{\delta})\,\sigma./\delta_1^2\right] + O(m^{-2}). \tag{3.17}$$

Turning to the evaluation of $E_2\phi_{2*}$, it follows from (3.13) that

$$E_2\phi_{2*} = (1 - 1/m)\left[\sigma.^2 + 2\sum_i \sigma_i^2/(n-1)m^2\right][1 + \delta_1^{-2}E_2(\hat{\delta} - \delta_1)^2]/n\delta_1,$$

where

$$E_2(\hat{\delta} - \delta_1)^2 = 2(m-1)^{-2}\left[(1 + 2/m)\sum_i \delta_{1i}^2 + {}_1^2\right].$$

Hence, neglecting terms of order less than m^{-1}, we obtain

$$E_2\phi_{2*} = (n\delta_1)^{-1}\left[\sigma.^2 + 2\sum_i \sigma_i^2/(n-1)m^2 - \sigma.^2/m + 2\sigma.^2\sum_i \delta_{1i}^2/m^2\delta_1^2\right] + O(m^{-3/2}). \tag{3.18}$$

We can now write, using (3.17) and (3.18),

$$\begin{aligned} E_2(\phi_{2*} - \phi_{1*}) = n\delta_1^{-1}\{&-(\delta_1 - \tau)^2 + m^{-1}[(3 + 2/(n-1))(\delta_1 - \tau)^2 \\ &+ 2s^2(\tilde{\delta})(n(n-1)^{-1} - (\tau/\delta_1)^2)]\} + O(m^{-3/2}). \end{aligned} \tag{3.19}$$

In arriving at (3.19), we used the following identities:

$$m^{-1}\sum_i \delta_{1i}^2 = s^2(\tilde{\delta}) + \delta_1^2, \qquad m^{-1}n^{-2}\sum_i \sigma_i^2 = s^2(\tilde{\delta}) + (\delta_1 - \tau)^2,$$

$$n^{-1}\sigma. = (\delta_1 - \tau), \qquad m^{-1}n^{-1}\sum_i \delta_{1i}\sigma_i = s^2(\tilde{\delta}) + \delta_1(\delta_1 - \tau).$$

3.2.4. *Expectation over* $\tilde{\sigma}$

Substituting (3.19) in (3.7), it remains to evaluate the expectation over $\tilde{\sigma}$ to arrive at a second-order approximation to MSE of the EBLUP. Expressing $E(\sigma./n) = \beta/n = E(\delta_1 - \tau)$, the formula (3.7) may be written as

$$\begin{aligned} E(\hat{\mu}_i - \mu_i)^2 = E[&\tau(1 - \tau/\delta_1) + m^{-1}\delta_1^{-1}\{(3 + 2/(n-1))(\delta_1 - \tau)^2/\delta_1 \\ &+ 2s^2(\tilde{\delta})[(n/(n-1)) - (\tau/\delta_1)^2]\}] + O(m^{-3/2}). \end{aligned} \tag{3.20}$$

Note that all terms on the right-hand side of (3.20) are nonnegative, since $\delta_1 - \tau \geqslant 0$ and $n/(n-1) \geqslant (\tau/\delta_1)^2$.

Now an approximation to the leading term, $1 - \tau E\delta_1^{-1}$, follows by expanding δ_1^{-1} as

$$\delta_1^{-1} = \delta^{-1}[1 - (\delta_1 - \delta)/\delta + (\delta_1 - \delta)^2/\delta^2] - \delta_1^{-1}(\delta_1 - \delta)^3/\delta^3, \quad (3.21)$$

where $\delta = E\delta_1 = \tau + \beta/n$. Since $\delta_1 > \tau$, the expectation of the remainder term in (3.21) is bounded by a constant times $E(\delta_1 - \delta)^3$ which is of order $O(m^{-3/2})$, noting that $\delta_1 - \delta = n^{-1}(\sigma . - \beta)$ and that the σ_i are i.i.d. random variables with $E\sigma_i = \beta$. Hence

$$E\delta_1^{-1} = \delta^{-1}[1 + \alpha/mn^2\delta^2] + O(m^{-3/2})$$

and

$$\tau E(1 - \tau/\delta_1) = \tau[\beta - \tau\alpha/nm\delta^2]/n\delta + O(m^{-3/2}). \quad (3.22)$$

Unfortunately the approximation (3.22) is not always positive, although the left-hand side of (3.22) is a nonnegative quantity. Therefore, in case a negative value is obtained, zero or the absolute value will serve as good approximation. A simple sufficient condition for $\beta - \tau\alpha/nm\delta^2 > 0$ is $m\beta^2 - \alpha > 0$ which we can easily take as an assumption, since our approximation is restricted to large m anyway. But note that some special distributions for $\tilde{\sigma}$, giving large weight to small error variances and still having large expectations due to large tails, will not satisfy this assumption for any m.

Turning now to the $O(m^{-1})$ terms in (3.20), we use the identity

$$\frac{f(\tilde{\delta})}{\delta_1} = \frac{f(\tilde{\delta})}{\delta}\left[1 - \frac{(\delta_1 - \delta)}{\delta}\frac{\delta}{\delta_1}\right] \quad (3.23)$$

for any function f of $\tilde{\delta} = (\tilde{\delta}_{11}, \ldots, \delta_{1m})'$. Assuming finite second moments for $f(\tilde{\delta})$, an application of Hölder's inequality to the remainder term of (3.23) gives

$$Ef(\tilde{\delta})/\delta_1 = Ef(\tilde{\delta})/\delta + O(m^{-1/2}). \quad (3.24)$$

It follows from (3.24) that

$$E(\delta_1 - \tau)^2/\delta_1 = (\delta - \tau)^2/\delta + O(m^{-1/2}) \quad (3.25)$$

and

$$Es^2(\tilde{\delta})/\delta_1 = \alpha/\delta n^2 + O(m^{-1/2}). \quad (3.26)$$

Finally, it follows from (3.23) that

$$\frac{f(\tilde{\delta})}{\delta_1^3}=\frac{f(\tilde{\delta})}{\delta^3}\left(1-\frac{(\delta_1-\delta)}{\delta}\frac{\delta}{\delta_1}\right)^3$$
$$=f(\tilde{\delta})\,\delta^{-3}[1-3(\delta_1-\delta)/\delta_1+3(\delta_1-\delta)^2/\delta_1^2-(\delta_1-\delta)^3/\delta_1^3]. \tag{3.27}$$

Using Hölder's inequality, it follows from (3.26) that the expectations of all but the first term are $O(m^{-1/2})$ or less. Hence,

$$Es^2(\tilde{\delta})/\delta_1^3=\alpha/\delta^3n^2+O(m^{-1/2}). \tag{3.28}$$

Using (3.22), (3.25), (3.26), and (3.28) in (3.20), we obtain the second-order approximation to MSE of the EBLUP as

$$E(\hat{\mu}_i-\mu_i)^2=\tau\beta/n\delta-\tau^2\alpha/mn^2\delta^3+\beta^2/m\delta n^2$$
$$+2(m\delta n^2)^{-1}[n(\beta^2+\alpha)/(n-1)-\alpha\tau^2/\delta^2]+O(m^{-3/2}). \tag{3.29}$$

The first and the third terms on the right-hand side of (3.29) form the MSE of the BLUP, while the second and the fourth terms represent the change of MSE due to estimating the parameters β and δ in the BLUP. For the special case of equal error variances, i.e., $\alpha=0$, (3.29) reduces to the approximation derived by Prasad and Rao [3].

4. An Approximately Unbiased Estimator of MSE

Looking at the expression (3.5) for the MSE of $\hat{\mu}_i$, we see that $(\hat{\beta}-\phi_1+\phi_2)/n$ is an unbiased estimator of MSE, provided that $e_{i\cdot}$ is known. Therefore, the estimation problem reduces to estimating $e_{i\cdot}$ and then evaluating the resulting bias. A natural estimator of $e_{i\cdot}$ is

$$\hat{e}_{i\cdot}=y_{i\cdot}-\hat{\mu}_i=\hat{c}(y_{i\cdot}-y_{\cdot\cdot}). \tag{4.1}$$

Substituting this estimator in the formula for ϕ_1, to define $\hat{\phi}_1$, we obtain a preliminary estimator of MSE as

$$\text{mse}_*(\hat{\mu}_i)=(\hat{\beta}-\hat{\phi}_1+\phi_2)/n=(\hat{\beta}-\phi_2)/n. \tag{4.2}$$

Its bias is given by

$$B=E\,\text{mse}_*(\hat{\mu}_i)-\text{MSE}\,(\mu_i)=E(\phi_1-2\phi_2)/n.$$

It remains to find an approximation to the estimator for B. Now noting that $E\phi_1 = EE_2\phi_{1*}$ and $E\phi_2 = EE_2\phi_{2*}$ and using the expressions (3.17) and (3.18) for $E_2\phi_{1*}$ and $E_2\phi_{2*}$, we obtain

$$E_2(\phi_1 - 2\phi_2) = 4(mn)^{-1}\left\{(\sigma./\delta_1)\left[\sigma. + s^2(\tilde{\delta})\,\sigma./\delta_1^2 - \sum_i \delta_{1i}\sigma_i/m\delta_1\right]\right.$$
$$\left. - \sum_i \sigma_i^2/m(n-1)\delta_1 - \sigma.^2 \sum_i \delta_{1i}^2/m\delta_1^3\right\} + O(m^{-3/2}).$$

The above expression can be further simplified as follows by using the identities below (3.19):

$$E_2(\phi_1 - 2\phi_2) = -4nm^{-1}\{s^2(\tilde{\delta})(\delta_1 - \tau)/\delta_1^2 + s^2(\tilde{\delta})/(n-1)\delta_1$$
$$+ n(\delta_1 - \tau)^2/(n-1)\delta_1\}. \tag{4.3}$$

Now using the expectations (3.25) and (3.26) in (4.3) and noting that

$$Es^2(\tilde{\delta})/\delta_1^2 = \alpha/(\delta n)^2 + O(m^{-1/2}), \tag{4.4}$$

we obtain

$$B = n^{-1}E(\phi_1 - 2\phi_2)$$
$$= -4(mn)^{-1}\{\beta^2/(n-1)\delta + \beta\alpha/n^2\delta^2 + \alpha/n(n-1)\delta\} + O(m^{-3/2}). \tag{4.5}$$

Hence, the order of the bias term, B, is $O(m^{-1})$. This suggests that we can correct the preliminary estimator (4.2) by estimating B, and the resulting estimator will be correct to terms of order m^{-1}.

We now turn to the estimation of bias B given by (4.5). To construct an estimator of β^2/δ, we observe that

$$E\hat{\beta}^2/\hat{\delta} = E\left[\sigma.^2 + 2\sum_i \sigma_i^2/m^2(n-1)\right]\delta_1^{-1}\left[1 + 2\sum_i \delta_{1i}^2/m^2\delta_1^2\right] + O(m^{-3/2}), \tag{4.6}$$

using (3.6) and

$$E_2\hat{\delta}^{-1} = \delta_1^{-1}\left(1 + 2\sum_i \delta_{1i}^2/m^2\delta_1^2\right) + O(m^{-3/2}).$$

It now follows from (4.6) that

$$E(\hat{\beta}^2/\hat{\delta}) = \beta^2/\delta + O(m^{-1}). \tag{4.7}$$

Turning to the estimation of α/δ, we consider the expression $m^{-1}\sum_i \hat{\sigma}_i^2$ with $\hat{\sigma}_i$ defined in (2.3) whose expectation is

$$Em^{-1}\sum_i \hat{\sigma}_i^2 = E(n+1)(n-1)^{-1}\sum_i \sigma_i^2/m = (n+1)(n-1)^{-1}(\beta^2+\alpha), \tag{4.8}$$

noting that $(n-1)\,\sigma_i^{-1}\hat{\sigma}_i = \sigma_i^{-1}\sum_j (y_{ij}-y_{i\cdot})^2$ is a χ^2 variable with $n-1$ degrees of freedom, given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$. It follows from (4.8) that

$$\hat{\alpha} = (n-1)(n+1)^{-1}\,m^{-1}\sum_i \hat{\sigma}_i^2 - \hat{\beta}^2 \tag{4.9}$$

is an estimator for α with bias of order m^{-1}. In fact,

$$\begin{aligned} E\hat{\alpha}/\hat{\delta} &= E\delta_1^{-1}\left(m^{-1}\sum_i \sigma_i^2 - \sigma_\cdot^2\right) + O(m^{-1/2}) \\ &= En^2 s^2(\tilde{\delta})/\delta_1 + O(m^{-1/2}) \\ &= \alpha/\delta + O(m^{-1/2}). \end{aligned} \tag{4.10}$$

The last step follows from (3.26). It follows from (4.10) that $\hat{\alpha}/\hat{\delta}$ estimates α/δ to the desired order of approximation.

Finally, to estimate $\alpha\beta/\delta^2$ we try $\hat{\alpha}\hat{\beta}/\hat{\delta}^2$. Routine computation, using cov $(\hat{\alpha}, \hat{\beta}) = O(m^{-1})$, yields

$$\begin{aligned} E\hat{\alpha}\hat{\beta}/\hat{\delta}^2 &= E\sigma_\cdot(\Sigma\sigma_i^2/m - \sigma_\cdot^2)/\delta_1^2 + O(m^{-1/2}) \\ &= En^3 s^2(\tilde{\delta})(\delta_1 - \tau)/\delta_1^2 + O(m^{-1/2}) \\ &= \alpha\beta/\delta^2 + O(m^{-1/2}). \end{aligned} \tag{4.11}$$

The last step follows from (3.26) and (4.4).

It now follows from (4.5), (4.7), (4.10), and (4.11) that an estimator of B to the desired order of approximation is given by

$$\hat{B} = -4[\hat{\beta}\hat{\alpha}/n\hat{\delta} + n\hat{\beta}^2/(n-1) + \hat{\alpha}/(n-1)]/mn^2\hat{\delta}, \tag{4.12}$$

i.e., $E\hat{B} = B + O(m^{-3/2})$. An estimator of MSE correct to terms of order m^{-1} is now obtained as

$$\begin{aligned} \mathrm{mse}(\hat{\mu}_i) &= \mathrm{mse}_*(\hat{\mu}_i) - \hat{B} \\ &= \hat{c}(\hat{\tau} + \hat{\beta}/nm) + 4(mn^2\hat{\delta})^{-1} \\ &\quad \times [(\hat{c} + (n-1)^{-1})(\hat{\alpha} + \hat{\beta}^2) + \hat{\tau}\hat{\beta}^2\hat{\delta}^{-1}], \end{aligned} \tag{4.13}$$

where $\hat{\tau} = \hat{\delta} - \hat{\beta}/n$ and $\hat{\alpha}$ and $\hat{c}$ are given by (4.9) and (2.4), respectively. Note that the first term in (4.13) is the naive estimator of MSE obtained by ignoring the uncertainty in the estimators $\hat{\tau}$ and $\hat{\beta}$ and using the MSE of the BLUP as an approximation to the true MSE of the EBLUP. This naive estimator,

$$\text{mse}_N(\hat{\mu}_i) = \hat{c}(\hat{\tau} + \hat{\beta}/nm), \tag{4.14}$$

could lead to a serious understatement; see Section 5.

5. Simulation Study

We performed a small simulation study to investigate the finite sample accuracy of the second-order MSE approximation (3.29), denoted by $\text{MSE}_A(\hat{\mu}_i)$, and the relative bias of $\text{mse}(\hat{\mu}_i)$, the approximately unbiased estimator of MSE of the EBLUP $\hat{\mu}_i$. To this end, we employed the following parameter values: $\mu = 0$, $\tau = 1$ (without loss of generality), $\beta = 5$, $m = 30$, and two values of n: 3 and 10. Using these parameter values, we generated 10,000 independent data sets $\{y_{i\cdot}, \hat{\sigma}_i; i = 1, ..., 30\}$ as follows:

Step 1. For each set, generate $\sigma_1, ..., \sigma_{30}$ from a χ^2 distribution with $\beta = 5$ degrees of freedom ($\alpha = 2\beta = 10$ in this case).

Step 2. For each set generate μ_i from $N(0, 1)$ and $e_{i\cdot}$ from $N(0, \sigma_i/n)$, $i = 1, ..., 30$. Let $y_{i\cdot} = \mu_i + e_{i\cdot}$, $i = 1, ..., 30$. Further generate a_i from a χ^2 distribution with $n-1$ degrees of freedom, and let $\hat{\sigma}_i = a_i \sigma_i/(n-1)$, $i = 1, ..., 30$. All the variables were generated independently from the specified distributions.

Without loss of generality, we consider the estimation of the true MSE of $\hat{\mu}_1$, since the sample size, n, is the same for all the m groups. The EBLUP $\hat{\mu}_1$, $\text{mse}(\hat{\mu}_1)$, and the naive estimator of MSE, $\text{mse}_N(\hat{\mu}_1)$, were computed from each data set $\{y_{i\cdot}, \hat{\sigma}_i; i = 1, ..., 30\}$. Simulated values of $\text{MSE}(\hat{\mu}_1)$, the true MSE of $\hat{\mu}_1$, $\text{Emse}(\hat{\mu}_1)$, and $\text{Emse}_N(\hat{\mu}_1)$ wer then computed from the 10,000 values of $\hat{\mu}_1$, $\text{mse}(\hat{\mu}_1)$, and $\text{mse}_N(\hat{\mu}_1)$ so generated. These values, along with the relative bias of $\text{mse}(\hat{\mu}_1)$ and $\text{mse}_N(\hat{\mu}_1)$, as estimators of the true MSE of $\hat{\mu}_1$, are reported in Table I. We also calculated the values of the second-order MSE approximation and the MSE of the BLUP $\tilde{\mu}_1$, using the specified parameter values. These values are also reported in Table I.

TABLE I

Simulated values of $\text{MSE}(\hat{\mu}_1)$, $\text{Emse}(\hat{\mu}_1)$, $\text{Emse}_N(\hat{\mu}_1)$ and percent relative biases of $\text{mse}(\hat{\mu}_1)$ and $\text{mse}_N(\hat{\mu}_1)$.

n	$\text{MSE}(\hat{\mu}_1)$	$\text{MSE}_A(\hat{\mu}_1)$	$\text{MSE}_N(\hat{\mu}_1)$	$\text{Emse}(\hat{\mu}_1)$	$\text{Emse}_N(\hat{\mu}_1)$	RB	RB_N
3	0.805	0.800	0.660	0.833	0.527	0.034	−0.346
10	0.355	0.353	0.339	0.356	0.324	0.002	−0.088

Note. $RB = [\text{Emse}(\hat{\mu}_1) - \text{MSE}(\hat{\mu}_1)]/\text{MSE}(\hat{\mu}_1)$, $RB_N = [\text{Emse}_N(\hat{\mu}_1) - \text{MSE}(\hat{\mu}_1)]/\text{MSE}(\hat{\mu}_1)$.

It is clear from Table I that the second-order MSE approximation is very accurate ever for small $n(=3)$: $\mathrm{MSE}_A(\hat{\mu}_1)=0.800$ compared to $\mathrm{MSE}(\hat{\mu}_1)=0.805$. On the other hand, the use of MSE of the BLUP, denoted by $\mathrm{MSE}_N(\hat{\mu}_1)$, as an approximation to the true MSE of the EBLUP, leads to serious understatement for $n=3$: $\mathrm{MSE}_N(\hat{\mu}_1)=0.660$ compared to $\mathrm{MSE}(\hat{\mu}_1)=0.805$. The understatement, however, is less serious for $n=10$: $\mathrm{MSE}_N(\hat{\mu}_1)=0.339$ compared to $\mathrm{MSE}(\hat{\mu}_1)=0.355$.

Turning to the relative bias of the estimators of MSE, we see from Table I that the relative bias of $\mathrm{mse}(\hat{\mu}_1)$ is very small even for $n=3$: 3.4%. On the other hand, the naive estimator, $\mathrm{mse}_N(\hat{\mu}_1)$, leads to serious underestimation for $n=3$, since its relative bias is -34.6%. The underestimation, however, is less serious for $n=10$, since the relative bias is reduced to -9%.

Our simulation study has confirmed the accuracy of the second-order MSE approximation and the approximate unbiasedness of the estimator of MSE, for large m.

6. Proofs

6.1. *Proofs of Lemmas* 1 *and* 2

Proof of Lemma 1. The estimator $\hat{\delta}$ may be written in matrix form as

$$\hat{\delta}=z'Mz/(m-1),$$

where $z=(y_{1\cdot},\ldots,y_{m\cdot})'$ and $M=I_m-1_m1'_m/m$ with I_m and 1_m denoting the identity matrix of order m and the vector of m unit elements, respectively. Further, we can write $M=BB'$, where B is a $m\times(m-1)$ matrix such that $B'B=I_{m-1}$. We form the random variable $X=z'B(B'DB)^{-1}B'z$ with $D=\mathrm{Diag}_i(\tau+\sigma_i/n)$ which has a χ^2 distribution on m $-$ 1 degrees of freedom. If the symbols $\alpha_{\min}(A)$ and $\alpha_{\max}(A)$ denote the smallest and largest eigenvalues of a matrix A, then

$$\alpha_{\max}[(B'DB)^{-1}]\leqslant[\alpha_{\min}(D)]^{-1}\leqslant\tau^{-1}.$$

Therefore, $X\leqslant z'BB'z/\tau=z'Mz/\tau$ and

$$\hat{\delta}^{-k}=[(m-1)/z'Mz]^k\leqslant[(m-1)/\tau]^k\,X^{-k}.$$

Also since X is χ^2 variable with $m-1$ degrees of freedom,

$$EX^{-k}=2^{-k}\prod_{i=1}^{k}\left(\frac{m-1}{2}-i\right)^{-1},$$

provided $m > 2k + 1$. Therefore,

$$E\hat{\delta}^{-k} \leqslant \tau^{-k} \prod_{i=1}^{k} \left(1 - \frac{2i}{m-1}\right)^{-1} \leqslant (2/\tau)^k. \quad \blacksquare$$

Proof of Lemma 2. By Hölder's inequality, we have

$$E[(\hat{\delta} - \delta_1)/\delta_1]^s f/\hat{\delta} \leqslant (Ef^2)^{1/2} [E(\hat{\delta} - \delta_1)^{4s} \delta_1^{-4s}]^{1/4} (E\hat{\delta}^{-4})^{1/4}. \quad (6.1)$$

The first term on the right-hand side of (6.1) is bounded by assumption, while the last term is bounded by Lemma 1 if $m > 17$. It remains to show that

$$E(\hat{\delta} - \delta_1)^{4s} \delta_1^{-4s} = O(m^{-2s}). \quad (6.2)$$

We first investigate the conditional expectation given $\tilde{\sigma} = (\sigma_1, ..., \sigma_m)'$. Using the result iv(c), page 39 in Rao and Kleffe [4], the conditional expectation satisfies

$$E_2(\hat{\delta} - \delta_1)^{4s} \delta_1^{-4s} \leqslant K[\text{tr } MDMD/\delta_1^2]^{2s}/(m-1)^{4s},$$

where M and D are as defined in the proof of Lemma 1, tr denotes the trace operator, and K is a constant independent of m. Also, by using $M \leqslant I$,

$$\text{tr } MDBD \leqslant \sum_{i=1}^{m} (\tau + \sigma_i/n)^2,$$

so that by Minkovski's inequality,

$$E_2(\hat{\delta} - \delta_1)^{4s} \delta_1^{-4s} \leqslant \frac{K}{(m-1)^{4s}} \left\{ \sum_{i=1}^{m} \left(E\left(\frac{\tau + \sigma_i/n}{\delta_1}\right)^{4s} \right)^{-2s} \right\}^{2s}.$$

The result (6.2) now follows from

$$E[(\tau + \sigma_i/n)/\delta_1]^{4s} \leqslant \tau^{-4s} E(\tau + \sigma_i/n)^{4s} = O(1),$$

since the σ_i have finite $4s$th moments.

6.2. *Derivation of* (3.15) *and* (3.16)

6.2.1. *Computation of* $E_2 \sum_i (y_i. - y..) e_i.\hat{\delta}$

We first express $(y_i. - y..)$, $e_i.$, and $\hat{\delta}$ as linear and quadratic forms in the random vectors $\mu = (\mu_1, ..., \mu_m)'$ and $e = (e_i., ..., e_m.)'$; we obtain

$$\sum_i (y_i. - y..) e_i. = e'M(\mu + e), \qquad \hat{\delta} = (\mu + e)' M(\mu + e)/(m-1).$$

Using now the independence of μ and e and $ME\mu=0$, we obtain

$$E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta} = E_2[e'Me\mu'M\mu + 2(e'M\mu)^2 + (e'Me)^2]/(m-1).$$

Next noting that μ and e have covariance matrices τI_m and $\mathrm{Diag}_i(\sigma_i/n)$, respectively, and using general results on moments of the above forms in normal variables (see [4, p. 53]), we obtain

$$E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta} = \left[(m-1)^2\, \sigma_{\cdot}\delta_1 + 2\sum_i \sigma_i \delta_{1i}\right] \Big/ n(m-1) + O(m^{-1})$$

which is equivalent to the expression (3.15).

6.2.2. *Computation of* $E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}^2$

Expressing the required expectation in terms of vectors μ and e yields

$$\begin{aligned} E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}^2 = E_2(\mu+e)'\, Me\{ & (\mu'M\mu)^2 + 4\mu'M\mu\mu'Me \\ & + 2\mu'M\mu e'Me + 4(\mu'Me)^2 \\ & + 4\mu'Mee'Me + (e'Me)^2\}/(m-1)^2. \end{aligned}$$

This expression simplifies because of the symmetry property of the multivariate normal distribution. Noting that the expectations of all order-3 products vanish, we arrive at

$$\begin{aligned} E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}^2 = E_2\{ & \mu'Me[4\mu'M\mu\mu'Me + 4\mu'Mee'Me] \\ & + e'Me[(\mu'M\mu)^2 + 2\mu'M\mu e'Me \\ & + 4(\mu'Me)^2 + (e'Me)^2\}/(m-1)^2. \end{aligned}$$

Computing term by term and neglecting terms of order less than m^{-1} leads to

$$\begin{aligned} E_2 \sum_i (y_{i\cdot}-y_{\cdot\cdot})\, e_{i\cdot}\hat{\delta}^2 = n^{-1} \Bigg\{ & (m+1)\, \sigma_{\cdot}\delta_1^2 + 4\delta_1 \sum_i \delta_{1i}\sigma_i/m \\ & + 2n^{-2}\sigma_{\cdot}\left(\sum_i \sigma_i^2/(m-1) - \sigma_{\cdot}^2\right)\Bigg\} + O(m^{-1}), \end{aligned}$$

an expression equivalent to (3.16).

ACKNOWLEDGMENTS

This work was supported by a research grant from the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the referees for their useful comments.

REFERENCES

[1] ARAGON, Y. (1984). Random variance linear models: estimation. *Comput. Statist. Quart.* **1** 295–309.

[2] GHOSH, M., AND MEEDEN, G. (1986). Empirical Bayes estimation in finite population sampling. *J. Amer. Statist. Assoc.* **81** 1058–1062.

[3] PRASAD, N. G. N., AND RAO, J. N. K. (1990). The estimation of the mean square error of small area estimators. *J. Amer. Statist. Assoc.* **55** 163–171.

[4] RAO, C. R., AND KLEFFE, J. (1988). *Estimation of Variance Components and Applications.* North-Holland, Amsterdam.